

What is claimed is:

1. An apparatus for synthesizing a singing voice of a song, comprising:

a storage section that stores template data in correspondence to various expressions applicable to music notes including an attack note and a non-attack note, the template data including first template data defining a temporal variation of a characteristic parameter for applying the corresponding expression to the attack note and second template data defining a temporal variation of a characteristic parameter for applying the corresponding expression to the non-attack note;

an input section that inputs voice information representing a sequence of vocal elements forming lyrics of the song and specifying expressions in correspondence to the respective vocal elements; and

a synthesizing section that synthesizes the singing voice of the lyrics from the sequence of the vocal elements based on the inputted voice information, such that the synthesizing section operates when the vocal element is of an attack note for retrieving the first template data corresponding to the expression specified to the vocal element and applying the specified expression to the vocal element of the attack note according to the retrieved first template data, and operates when the vocal element is of a non-attack note for retrieving the second template data

corresponding to the expression specified to the vocal element and applying the specified expression to the vocal element of the non-attack note according to the retrieved second template data.

2. The apparatus according to claim 1, wherein the synthesizing section includes a discriminating subsection that discriminates each vocal element to either of the non-attack note or the attack note based on the inputted voice information in real time basis during the course of synthesizing the singing voice of the song.

3. The apparatus according to claim 1, wherein the input section inputs the voice information containing timing information which specifies utterance timings of the respective vocal elements along a progression of the song, and wherein the synthesizing section includes a discriminating subsection that discriminates the respective vocal elements to either of the non-attack note or the attack note based on the utterance timings of the respective vocal elements, such that the vocal element is identified to the non-attack note when the vocal element has a preceding vocal element which is uttered before the vocal element and when a difference of the utterance timings between the vocal element and the preceding vocal element is within a predetermined time length, and otherwise the vocal element is identified to the attack

note when the vocal element has no preceding vocal element or has a preceding vocal element but the difference of utterance timings between the vocal element and the preceding vocal element exceeds the predetermined time length.

4. The apparatus according to claim 3, wherein the input section inputs the voice information in the form of a vocal element track and an expression track, the vocal element track recording the vocal elements integrally with the timing information such that the respective vocal elements are sequentially arranged along the vocal element track in a temporal order determined by the respective utterance timings, the expression track recording the expressions corresponding to the vocal elements in synchronization with the vocal element track.

5 The apparatus according to claim 1, wherein the input section inputs the voice information containing pitch information which represents a transition of a pitch applied to each vocal element in association with an utterance timing of each vocal element, and wherein the synthesizing section includes a discriminating subsection that discriminates each vocal element to either of the non-attack note or the attack note based on the pitch information, such that the vocal element is identified to the non-attack note when a value of the pitch is found in a

preceding time slot extending back from the utterance timing of the vocal element by a predetermined time length, and otherwise the vocal element is identified to the attack note when a value of the pitch lacks in the preceding time slot.

6. The apparatus according to claim 5, wherein the input section inputs the voice information in the form of a vocal element track, a pitch track and an expression track, the vocal element track recording the sequence of the respective vocal elements in a temporal order determined by the respective utterance timings, the pitch track recording the transition of the pitch applied to each vocal element in synchronization with the vocal element track, the expression track recording the expressions corresponding to the vocal elements in synchronization with the vocal element track.

7. A program executable by a computer for synthesizing a singing voice of a song with use of template data stored in a storage correspondingly to various expressions applicable to music notes including an attack note and a non-attack note, the template data including first template data defining a temporal variation of a characteristic parameter for applying the corresponding expression to the attack note and second template data defining a temporal variation of a characteristic parameter for applying the

corresponding expression to the non-attack note, the program comprising:

a step of inputting voice information which represents a sequence of vocal elements forming lyrics of the song and which specifies expressions in correspondence to the respective vocal elements; and

a step of synthesizing the singing voice of the lyrics from the sequence of the vocal elements based on the inputted voice information, the step of synthesizing being carried out when the vocal element is of an attack note such that the first template data corresponding to the expression specified to the vocal element is retrieved from the storage and the specified expression is applied to the vocal element of the attack note according to the retrieved first template data, and being carried out when the vocal element is of a non-attack note such that the second template data corresponding to the expression specified to the vocal element is retrieved from the storage and the specified expression is applied to the vocal element of the non-attack note according to the retrieved second template data.

8. The program according to claim 7, wherein the step of synthesizing includes a sub step of discriminating each vocal element to either of the non-attack note or the attack note based on the inputted voice information in real time basis during the course of synthesizing the singing

voice of the song.

9. The program according to claim 7, wherein the step of inputting inputs the voice information containing timing information which specifies utterance timings of the respective vocal elements along a progression of the song, and wherein the step of synthesizing includes a sub step of discriminating the respective vocal elements to either of the non-attack note or the attack note based on the utterance timings of the respective vocal elements, such that the vocal element is identified to the non-attack note when the vocal element has a preceding vocal element which is uttered before the vocal element and when a difference of the utterance timings between the vocal element and the preceding vocal element is within a predetermined time length, and otherwise the vocal element is identified to the attack note when the vocal element has no preceding vocal element or has a preceding vocal element but the difference of utterance timings between the vocal element and the preceding vocal element exceeds the predetermined time length.

10. The program according to claim 9, wherein the step of inputting inputs the voice information in the form of a vocal element track and an expression track, the vocal element track recording the vocal elements integrally with the timing information such that the respective vocal

elements are sequentially arranged along the vocal element track in a temporal order determined by the respective utterance timings, the expression track recording the expressions corresponding to the vocal elements in synchronization with the vocal element track.

11 The program according to claim 7, wherein the step of inputting inputs the voice information containing pitch information which represents a transition of a pitch applied to each vocal element in association with an utterance timing of each vocal element, and wherein the step of synthesizing includes a sub step of discriminating each vocal element to either of the non-attack note or the attack note based on the pitch information, such that the vocal element is identified to the non-attack note when a value of the pitch is found in a preceding time slot extending back from the utterance timing of the vocal element by a predetermined time length, and otherwise the vocal element is identified to the attack note when a value of the pitch lacks in the preceding time slot.

12. The program according to claim 11, wherein the step of inputting inputs the voice information in the form of a vocal element track, a pitch track and an expression track, the vocal element track recording the sequence of the respective vocal elements in a temporal order determined by the respective utterance timings, the pitch track recording

the transition of the pitch applied to each vocal element in synchronization with the vocal element track, the expression track recording the expressions corresponding to the vocal elements in synchronization with the vocal element track.

13. A method of synthesizing a singing voice of a song, comprising:

a step of storing template data in a storage correspondingly to various expressions applicable to music notes including an attack note and a non-attack note, the template data including first template data defining a temporal variation of a characteristic parameter for applying the corresponding expression to the attack note and second template data defining a temporal variation of a characteristic parameter for applying the corresponding expression to the non-attack note;

a step of inputting voice information which represents a sequence of vocal elements forming lyrics of the song and which specifies expressions in correspondence to the respective vocal elements; and

a step of synthesizing the singing voice of the lyrics from the sequence of the vocal elements based on the inputted voice information, the step of synthesizing being carried out when the vocal element is of an attack note such that the first template data corresponding to the expression specified to the vocal element is retrieved from



the storage and the specified expression is applied to the vocal element of the attack note according to the retrieved first template data, and being carried out when the vocal element is of a non-attack note such that the second template data corresponding to the expression specified to the vocal element is retrieved from the storage and the specified expression is applied to the vocal element of the non-attack note according to the retrieved second template data.